The Logic of Risky Knowledge

Henry E. Kyburg, Jr. [1,2]

Philosophy and Computer Science, University of Rochester,
Rochester NY 14627, USA
and
Institute for Human and Machine Cognition, University of West Florida,
40 South Alcaniz Street, Pensacola FL 32501, USA

Choh Man Teng [1,3]

Institute for Human and Machine Cognition, University of West Florida,
40 South Alcaniz Street, Pensacola FL 32501, USA

Abstract

Much of our everyday knowledge is risky. This not only includes personal judgments, but the results of measurement, data obtained from references or by report, the results of statistical testing, etc. There are two (often opposed) views in AI on how to handle risky empirical knowledge. One view, characterized often by modal or nonmonotonic logics, is that the structure of such knowledge should be captured by the formal logical properties of a set of sentences, if we can just get the logic right. The other view takes probability to be central to the characterization of risky knowledge, but often does not allow for the tentative or corrigible acceptance of a set of sentences. We examine a view, based on ϵ-acceptability that combines both probability and modality. A statement is ϵ-accepted if the probability of its denial is at most ϵ, where ϵ is taken to be a fixed small parameter as is customary in the practice of statistical testing. We show that given a body of evidence Γ_δ, the set of ϵ-accepted statements Γ_ϵ has exactly the logical structure of a classical modal system **EMN**, the smallest classical modal logic **E** supplemented by the schemata **M**: $\Box_\epsilon(\ulcorner \phi \wedge \psi \urcorner) \rightarrow (\Box_\epsilon \phi \wedge \Box_\epsilon \psi)$ and **N**: $\Box_\epsilon \top$.

Key words: risky knowledge, ϵ-acceptability, statistical testing

[1] This work has been supported in part by the National Science Foundation grants STS-9906128 and ITS-0082928, and NASA grant NCC2-1239.

[2] Email: `hkyburg@ai.uwf.edu`

[3] Email: `cmteng@ai.uwf.edu`

1 Introduction

Certainly, a lot of what we consider *knowledge* is risky. "Do you know when the next plane to L.A. leaves?" "Yes; 11:04; I just looked it up." You know that it is raining; you just looked out the window and saw that it was. John knows the length of the table; he just measured it and knows that it is 42.0 ± 0.10 in. We know that of the next thousand births in this big hospital, at least 400 will be male. Since we just obtained evidence in the 0.01 rejection region, we know that H_0 is false.

In general, particularly in the case of measurement, it seems unreasonable to deny that statements like these, despite the fact that evidence can render them no more than highly probable, can qualify as knowledge.

What is the logical structure of the set of conclusions that may be obtained by nonmonotonic or inductive inference from a body of evidence? Is it a deductively closed theory, something weaker, something stronger? We make minimal assumptions about evidence, and take inductive inference to be based on probability thresholding [11]. We show that the structure of the set of inferred statements corresponds to a minimal modal logic, in which the necessity operator is interpreted as "it is known that." This logic is not closed under conjunction, but is otherwise normal.

Of course one may encounter the objection: "Well, you don't really *know*...; it is merely probable." But this is ingenuous or quibbling, *unless there are specific grounds* for assigning a lower probability than that initially suggested.

We are interested in the *objective* risk of error. Our assessment of risk is to be based on evidence, which in fact will include statistical knowledge (as when we accept as "known" the approximate results of measurement). For reasons discussed elsewhere [5], we must make a sharp distinction between the set of sentences constituting the evidence, and the set of nonmonotonically or inductively inferred sentences. The evidence itself may be uncertain, but we shall suppose that it carries less risk of error than the risky knowledge we derive from it. Let us take the set of sentences constituting the *evidence* to be Γ_δ and the set of sentences constituting our risky knowledge Γ_ϵ, where $\delta \leq \epsilon$.

The set of sentences Γ_δ may even have the same structure as Γ_ϵ, and in turn may be derived from a set of less risky evidence statements Γ_γ. Such a structure would *allow* the question of inductive or uncertain support or justification to be raised at any level, but would not, of course *require* it.

We shall suppose that Γ_δ and Γ_ϵ are sets of sentences of a general first order language that includes means of referring to real numbers. In addition the language will include statements whose interpretation is that "the proportion of objects $r(x)$ that are also $t(x)$ lies between p and q." Finally, we shall assume that Γ_δ, the set of evidential statements, is not empty, and that it contains logical and mathematical truths.

Note that the small number "ϵ" is to be construed as a fixed number, rather than a variable that approaches a limit. Both Ernest Adams [1] and

Judea Pearl [8] have sought to make a connection between high probability and modal logic, but both have taken probabilities *arbitrarily* close to one as corresponding to knowledge. In real life we do not have access to probabilities arbitrarily close to one. Thus we have chosen to follow the model of hypothesis testing in statistics: we reject a hypothesis (accept its complement) when the chance of doing so erroneously is less than a fixed finite amount.

2 The Sentences of Γ_ϵ

We want our risky knowledge to be objectively justified by the evidence. We shall interpret that to mean that the *probability*, relative to Γ_δ, that a statement S in Γ_ϵ is false is to be no more than a fixed value ϵ. We shall explain loosely what we mean by "probability," state a number of its properties, and then prove a number of theorems about the acceptability of statements in Γ_ϵ.

2.1 Probability

We will follow the treatment of probability in [6]. On this approach probability is interval-valued, defined in a (two-sorted) first order logic, relativized to evidence, and based on known statistical frequencies. For example, given that we know that between 70% and 80% of the balls in an urn are black, and that we know nothing about the next ball to be drawn that indicates that it is special in any way with respect to color, we would say that the probability of the sentence "The next ball to be drawn is black," relative to what we know, is [0.7,0.8]. Knowing that errors of measurement of length by method M are normally distributed $N(0.0, 0.02)$, and that a measurement of a yielded the value 11.31, we can be 0.95 confident that the length of a lies between 11.27 and 11.35, i.e., the probability of "$11.27 \leq length(a) \leq 11.35$," is [0.95,0.95]. A test of a hypothesis of size α that yields a point in the rejection region supports the denial of the null hypothesis H_0 to the degree $[1-\alpha, 1]$, or runs a risk of error of at most α.

We represent the probability of the statement S, given the background knowledge Γ_δ by $\mathrm{Prob}(S, \Gamma_\delta)$. There are a number of important facts about probability as we construe it, that we will simply list here:

(i) Given a body of evidence Γ_δ, every statement S of our language has a probability: there are p and q such that $\mathrm{Prob}(S, \Gamma_\delta) = [p, q]$.

(ii) Probability is unique: If $\mathrm{Prob}(S, \Gamma_\delta) = [p, q]$ and $\mathrm{Prob}(S, \Gamma_\delta) = [r, s]$ then $p = r$ and $q = s$.

(iii) If S and T are known to have the same truth value, i.e., if the biconditional $\ulcorner S \equiv T \urcorner$ is in Γ_δ, they have the same probability: If $\ulcorner S \equiv T \urcorner \in \Gamma_\delta$, then $\mathrm{Prob}(S, \Gamma_\delta) = \mathrm{Prob}(T, \Gamma_\delta)$. [4]

[4] We follow Quine [9] in using quasi-quotation (corners) to specify the *forms* of expressions in our formal language. This $\ulcorner S \equiv T \urcorner$ becomes a specific biconditional expression on the

(iv) If $\text{Prob}(S, \Gamma_\delta) = [p, q]$, then $\text{Prob}(\ulcorner \neg S \urcorner, \Gamma_\delta) = [1 - q, 1 - p]$. The probability of the negation of a statement is determined by the probability of the statement itself.

(v) If S entails T, i.e., if $S \vdash T$, $\text{Prob}(S, \Gamma_\delta) = [p_S, q_S]$ and $\text{Prob}(T, \Gamma_\delta) = [p_T, q_T]$ then $p_S \leq p_T$ and $q_S \leq q_T$.

(vi) If $S \in \Gamma_\delta$, then $\text{Prob}(S, \Gamma_\delta) = [1, 1]$; If $\ulcorner \neg S \urcorner \in \Gamma_\delta$, then $\text{Prob}(S, \Gamma_\delta) = [0, 0]$.

(vii) The semantics underlying probability reflect known relative frequencies, though the actual probability interval of a statement may be the cover of several frequency intervals.

2.2 The Syntax of ϵ-acceptability

It is natural to suggest that it is worth accepting a statement S as known if there is only a negligible chance that it is wrong. Put in terms of probability, we might say that it is reasonable to accept a statement in Γ_ϵ when the maximum probability of its denial relative to what we take as evidence, Γ_δ, is less than or equal to ϵ. This suggests the following definition of Γ_ϵ, our body of risky knowledge, in terms of our evidence Γ_δ:

Definition 2.1 $\Gamma_\epsilon = \{S : \exists p, q \; (\text{Prob}(\ulcorner \neg S \urcorner, \Gamma_\delta) = [p, q] \wedge q \leq \epsilon)\}$.

We stipulate that a sentence belongs to the set of statements accepted at level ϵ when the *maximum* chance of its being wrong (q) is less than or equal to our tolerance for error, ϵ. This reflects — and is in part motivated by — the theory of testing statistical hypotheses.

A few easy theorems will establish some important facts about the structure of Γ_ϵ. First we restate Definition 2.1 in a more positive form, making use of property (iv) of probability in Section 2.1:

Theorem 2.2 $S \in \Gamma_\epsilon \equiv \exists p, q \; (\text{Prob}(S, \Gamma_\delta) = [p, q] \wedge p \geq 1 - \epsilon)$.

An important fact concerning the use of inference in Γ_ϵ is captured by the following theorem:

Theorem 2.3 *If $S \in \Gamma_\epsilon$ and $S \vdash T$ and $S \vdash T'$ then $T \wedge T' \in \Gamma_\epsilon$.*

Proof. This follows from property (v) of probability in Section 2.1, together with the principle that if $S \vdash T$ and $S \vdash T'$, then $S \vdash T \wedge T'$. □

But we do not have adjunction in general:

Theorem 2.4 *It is possible that $S \in \Gamma_\epsilon$ and $T \in \Gamma_\epsilon$ but $\ulcorner S \wedge T \urcorner \notin \Gamma_\epsilon$.*

Proof. The probability of S and of T may each exceed $1 - \epsilon$, while the probability of their conjunction does not. □

---

replacement of S and T by specific formulas of the language.

The fact that we do not in general have adjunction may strike some people as disastrous. After all, adjunction is a basic form of inference. In mitigation, we point out that Theorem 2.3 shows that there are cases in which we clearly *do* have adjunction. What is required is that the items that are to be adjoined be derivable from a single statement that is itself acceptable.

When do we call on adjunction? When we have an argument that proceeds from a *list* of premises to a conclusion. This is convenient, and may be perspicuous, but is not essential; the same conclusion could be derived from the conjunction of the premises without using adjunction. Where it makes a difference is where the premises are *individually* supported by empirical evidence. But in this case the persuasiveness of the argument depends in part on the empirical support given to the conjunction of the premises. Too many premises, each only *just* acceptable, will not provide good support for a conclusion even when the conclusion does validly follow from the premises.

For example, from "Cow #1 shows no unusual symptoms" and "If a cow shows no unusual symptoms then she will have a normal calf," one may infer "Cow #1 will have a normal calf." From a thousand sets of premises of that form, one may validly infer (making use of adjunction) that all thousand of the cows will have normal calves. But while the inference is valid, the conclusion is not one that should be believed. Why should the conclusion not be believed? Because while the generalization will *almost* always be upheld by events, and so should be accepted, it will on rare occasions lead us astray.

Furthermore, note that although the conclusion about the thousand cows should not be believed, there is no particular premise that should be rejected. For a more complete discussion of adjunction and its role in logic, see [4].

To the extent that we are thinking of an argument as *supporting* its conclusion, any argument requires simultaneous (adjunctive) acceptance of its premises. Any doubt that infects the conjunction of the premises rightly throws a shadow on the conclusion. Except for notational convenience, any argument may be taken to have a single premise.

Many more people seem to have doubts about deductive closure than have doubts about conjunction. It is worth noting the connection between adjunction and deductive closure expressed by the following theorem:

Theorem 2.5 *If Σ is not empty and contains the first order consequences of any statement in it, then Σ is closed under conjunction if and only if Σ is deductively closed.*

3 Minimal Modal Logics

Following up an idea of Richard Montague and Dana Scott [10,7], Brian Chellas [3] fleshed out the idea of a "minimal modal logic." Since then, Arlo-Costa [2] has arrived at some interesting results concerning Barcan formulas in a first order modal logic, using the same models. We will follow Arlo-Costa in referring to the models underlying these logics as "neighborhood models." Our concerns here however will be restricted to the propositional case.

Let us first describe the terminology for minimal modal logics. We will denote by $\Box$ and $\Diamond$ the necessity and possibility operators, and $\mathcal{P}$ the set of propositional constants.

Definition 3.1 A tuple $\mathcal{M} = \langle W, N, \models \rangle$ is a *neighborhood model* iff

(i) W is a set [*worlds*].

(ii) $N : W \rightarrow 2^{2^W}$ is a function from W to sets of sets of worlds [*neighborhood function*].

(iii) $\models$ is a function $W \times \mathcal{P} \rightarrow \{0, 1\}$ [*truth assignment function*].

In this formulation, sentences are identified with sets of worlds. A sentence is represented by the maximal set of worlds in which it is true. A set of sentences is thus represented as a set of sets of worlds.

Informally, in a neighborhood model, for each world $w \in W$, there is an associated neighborhood $N(w)$ (given by the neighborhood function N) that consists of a set of sentences. The neighborhood of a world contains the set of sentences that are "necessary" at that world.

We will extend $\models$ to denote the valuation function of the model $\mathcal{M}$, and write $w \models_{\mathcal{M}} \phi$ if the sentence ϕ is true at the world w. In addition to the usual rules governing the truth of compound formulas, we have the following regarding the modal formulas.

Definition 3.2 Given a model $\mathcal{M} = \langle W, N, \models \rangle$ and a formula ϕ, the *truth set* of ϕ in $\mathcal{M}$, denoted by $||\phi||^{\mathcal{M}}$, is given by

$$||\phi||^{\mathcal{M}} = \{w \in W : w \models_{\mathcal{M}} \phi\}.$$

The truth of modal formulas at a world $w \in W$ is defined as follows.

$$w \models_{\mathcal{M}} \Box\phi \text{ iff } ||\phi||^{\mathcal{M}} \in N(w)$$
$$w \models_{\mathcal{M}} \Diamond\phi \text{ iff } ||\neg\phi||^{\mathcal{M}} \notin N(w)$$

The truth set of a formula contains all the worlds in which it is true. Recall that the neighborhood $N(w)$ of the world w contains all the sentences (represented by sets of worlds) that are necessary at w. Thus, the modal formula $\Box\phi$ is true at w iff the truth set of ϕ is contained in $N(w)$.

Similarly, the formula $\Diamond\phi$ is true at w iff the truth set of the negation of ϕ is *not* contained in $N(w)$. This corresponds to the conventional notion that the two modal operators are interdefinable: $\Diamond \equiv \neg\Box\neg$.

4 Knowledge as ϵ-acceptability

So far we have been calling the two modal operators the necessity and possibility operators. However, there can be other interpretations of these operators; they can be construed epistemically or deontically. For example, $\Box$ may be identified with the notion of belief, obligation, or requirements. The interpretation we are interested in is knowledge construed as ϵ-acceptability. In this case, $\Box\phi$ is interpreted as that ϕ is "known". A bit more precisely, given a set of evidence Γ_δ and a threshold ϵ, $\Box_\epsilon\phi$ corresponds to $\phi \in \Gamma_\epsilon$, and $\Diamond_\epsilon\phi$ corresponds to $\neg\phi \notin \Gamma_\epsilon$. In other words,

Definition 4.1

$\Box_\epsilon\phi : \mathrm{Prob}(\phi, \Gamma_\delta) = [p, q]$ and $p \geq 1 - \epsilon$;
$\Diamond_\epsilon\phi : \mathrm{Prob}(\phi, \Gamma_\delta) = [p, q]$ and $q > \epsilon$.

4.1 Constraints on the Neighborhood

Consider the axiom schema

($\Diamond$) $\Diamond\phi \leftrightarrow \neg\Box\neg\phi$,

and the rule of inference

(RE) $\dfrac{\phi \leftrightarrow \psi}{\Box\phi \leftrightarrow \Box\psi}$.

A modal system is called *classical* iff it contains the axiom schema **($\Diamond$)** and is closed under the rule of inference **(RE)**. The modal system **E**, consisting of **($\Diamond$)** and **(RE)**, is the smallest classical modal logic.

Theorem 4.2 *$\Box_\epsilon$ is classical.*

Proof. This follows from definition 4.1, and property (iii) of probability in Section 2.1. $\Box$

There are three conditions on neighborhood models that are of interest to us. Given a neighborhood model $\mathcal{M} = \langle W, N, \models \rangle$, for any world $w \in W$ and formulas ϕ and ψ, let us define the following conditions (we follow Arlo-Costa [2]).

(m) If $\ulcorner \phi \wedge \psi \urcorner \in N(w)$, then $\phi \in N(w)$ and $\psi \in N(w)$.
(c) If $\phi \in N(w)$ and $\psi \in N(w)$, then $\ulcorner \phi \wedge \psi \urcorner \in N(w)$.
(n) $W \in N(w)$.

These conditions correspond to the following schemata. (The symbol $\top$ denotes the truth constant.)

(M) $\Box_\epsilon(\ulcorner \phi \wedge \psi \urcorner) \rightarrow (\Box_\epsilon\phi \wedge \Box_\epsilon\psi)$.
(C) $(\Box_\epsilon\phi \wedge \Box_\epsilon\psi) \rightarrow \Box_\epsilon(\ulcorner \phi \wedge \psi \urcorner)$.
(N) $\Box_\epsilon\top$.

That is, the schema **(M)** is valid in the class of neighborhood models that satisfy **(m)**, and similarly for **(C)** and **(N)**.

4.2 The Modal System for ϵ-Acceptability

Recall that **E** is the smallest classical modal logic. A family of distinct classical minimal logics can be obtained by adopting in addition different combinations of the three axiom schemas **(M)**, **(C)**, and **(N)**. We will denote such systems by their lists of axioms. For example, **EMC** is the classical system containing the schemas **(M)** and **(C)**.

Among this family of modal logics, the one that is of particular interest to us is the system **EMN**. The following theorem establishes a link between **EMN**, a classical minimal logic, and Γ_ϵ, a system of ϵ-acceptability.

Theorem 4.3 *Given a body of evidence* Γ_δ *and a threshold* ϵ*,* $\Box_\epsilon$ *satisfies the modal system* **EMN**.

Proof. **EMN** is the minimal modal system with the axiom schemas **(**$\Diamond$**)**, **(M)**, and **(N)**, and a single rule of inference **(RE)**. By Theorem 4.2, $\Box_\epsilon$ is classical. Property (v) of probability in Section 2.1 yields **M** and **N**, provided Γ_δ is not empty. $\Box$

Note however, the axiom schema **(C)** is not valid in Γ_ϵ. Intuitively this corresponds to the rejection of adjunction. The conjunction of two ϵ-accepted formulas may itself have a probability that is not high enough to warrant admittance into Γ_ϵ, as mentioned in Theorem 2.4.

5 Conclusion

What is interesting about the results we have just presented is that they provide a reconciliation between an approach to knowledge in terms of probabilistic acceptance, and an approach to knowledge in terms of modal epistemic logics. We have shown that the same structure emerges when viewed in each way.

We have taken the set of evidence statements, Γ_δ, for granted. Presumably, this set of statements can be construed as risky, too. Note that, unless one is seeking "'ultimate" justification (whatever that may be) there is no dangerous circularity in this procedure. Given any degree of riskiness ϵ, we can ask for and receive objective justification of the statements in Γ_ϵ, in terms of a less risky set of statements Γ_δ. Γ_δ in turn can be questioned; this amounts to treating Γ_δ as derived from some even less risky set of statements Γ_η. And the same can be asked of Γ_η. That this can always be done provides us with all the non-circular objectivity we need. That it does provide us with objectivity is a consequence of the objectivity of probability as we construe it in [6].

In practise, of course, we can be perfectly satisfied with a body of evidence Γ_δ of some specified degree of maximal riskiness. What concerns us in practise is the question of whether that evidence provides adequate support for sentences in Γ_ϵ, the corpus of *practical* certainties we use for making decisions and calculating expectations. Many of these practical certainties will be the relative frequency statements we need for grounding probabilities. Again, there need be no circularity.

References

[1] Adams, E. W., "The Logic of Conditionals," Reidel, Dordrecht,, 1975.

[2] Arlo Costa, H., *First order extensions of classical systems of modal logic*, Studia Logica (forthcoming).

[3] Chellas, B., "Modal Logic," Cambridge University Press, 1980.

[4] Kyburg, H. E., Jr., *The rule of adjunction and reasonable inference*, Journal of Philosophy **94** (1997), pp. 109–125.

[5] Kyburg, H. E., Jr., *Real logic is nonmonotonic*, Minds and Machines **11** (2001), pp. 577–595.

[6] Kyburg, H. E., Jr. and C. M. Teng, "Uncertain Inference," Cambridge University Press, New York, 2001.

[7] Montague, R., *Universal grammar*, Theoria **36** (1970), pp. 373–398.

[8] Pearl, J., "Probabilistic Reasoning in Intelligent Systems," Morgan Kaufmann, San Mateo, 1988.

[9] Quine, W. V. O., "Mathematical Logic," Harvard, Cambridge, 1941.

[10] Scott, D., *Advice in modal logic*, in: K. Lambert, editor, *Philosophical Problems in Logic*, Reidel, Dordrecht, 1970 pp. 143–173.

[11] Teng, C. M., *Sequential threshholds: Context sensitive default extensions*, in: *Uncertainty in Artificial Intelligence 13* (1997), pp. 437–444.